Detection of Seizure EEG Signals Based on Reconstructed Phase Space of Rhythms in EWT Domain and Genetic Algorithm

Hesam Akbari$^1$, Somayeh Saraf Esmaili$^2*$ Sima Farzollah Zadeh$^3$

$^1$ Department of Biomedical Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran.
$^2$ Department of Biomedical Engineering, Garmsar Branch, Islamic Azad University, Garmsar, Iran.
$^3$ Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

Received: 15-Jan-2020, Revised: 29-Feb-2020, Accepted: 07-Mar-2020.

Abstract
Epilepsy is a brain disorder which stems from the abnormal activity of neurons and recording of seizures has primary interest in the evaluation of epileptic patients. A seizure is the phenomenon of rhythmicity discharge from either a local area or the whole brain and the individual behavior usually lasts from seconds to minutes. In this work, empirical wavelet transform (EWT) is applied to decompose signals into Electroencephalography (EEG) rhythms. EEG signals are separated into the delta, theta, alpha, beta and gamma rhythms using EWT. The proposed method has been evaluated by the benchmark dataset which is freely downloadable from the Bonn University website. Ellipse area (A) and shortest distance to 45 and 135-degree lines are computed from the 2D projection of reconstructed phase space (RPS) of rhythms as features. After that, the genetic algorithm is used as feature selection. Finally, selected features are fed to the K-nearest neighbor (KNN) classifier for the detection of the seizure (S) and seizure-free (SF) EEG signals. Our proposed method archived 98.33% accuracy in the classification of S and SF EEG signals with a tenfold cross-validation strategy that is higher than previous techniques.

Keywords: Electroencephalogram (EEG) Signals, Empirical Wavelet Transform (EWT), Reconstructed Phase Space (RPS), Genetic Algorithm, K-Nearest Neighbor (KNN) Classifier.

*Corresponding Authors Email: s.sesmaeily@iau-garmsar.ac.ir
1. INTRODUCTION

Epilepsy is a brain disorder caused by the abnormal activity of neurons. About 50 million people suffer from epilepsy. Most of whom live in developing countries [1]. One of the commonly used electrophysiological monitoring methods to detect epilepsy is Electroencephalography (EEG). Epileptic seizures in the human brain frequently manifest spikes in EEG signals which can be analyzed visually by the experts [2]. Visual inspection of long EEG records to detect the presence of epileptic seizures can be a cumbersome and time-consuming activity. Therefore, an automatic method to detect and classify the normal and epilepsy seizures is desirable. Recently, many methods have been developed to fulfill these proposes. Classification of the seizure (S) and seizure-free (SF) epilepsy EEG signals using permutation entropy have been reported in [3]. In [4], the Mean degree and the mean strength of horizontal visibility graph (HVG) have been used in the KNN classifier for detecting S EEG signals. Classification of S and SF EEG signals using the clustering technique and support vector machine (SVM) have been reported in [5]. In [6], the linear prediction error energy feature has been used for the classification of S and SF EEG signals. In [7], fractional linear prediction error (FLP) and signal energy are used as features for classification S and SF EEG signals. Empirical mode decomposition (EMD) has been proposed to decompose an input signal into intrinsic mode functions (IMFs) [8,9]. In [10], 95% confident area measure of second-order difference plot (SODP) from IMFs is extracted and an artificial neural network (ANN) classifier is applied to classify EEG signals in S and SF groups. In [11], dual-tree complex wavelet transform (DTCWT) decomposed EEG signals and various entropies and statistically based features computed as input to the general regression neural network (GRNN) classifier to discriminate S from SF EEG signals. Also, Classification of S and SF EEG signals using tunable-Q wavelet transform (TQWT) and Kraskov entropy have been reported in [12]. Recently, empirical wavelet transform (EWT) has been proposed to analyze non-stationary signals [13]. It decomposes the input signal based on an adaptive filter bank. Bandpass of EWT filters is determined using significant segmentation of the spectrum of the input signal.

In [14], the authors showed that EWT can decompose the EEG rhythms faster and more accurate than traditional wavelet transform. For this reason, in the present work, EWT is applied to decompose signals into EEG rhythms. Traditional wavelet transform and EMD separate these rhythms by multiplying number decompositions, while EWT can extract these rhythms by one-step proses. The reconstructed phase space (RPS) can manifest the dynamic of chaotic. Due to the chaotic nature of the EEG signals, RPS is applied to exhibit of EEG rhythms in 2D projection [15]. In this work, the ellipse area (A) and shortest distance to 45 and 135-degree lines are computed from the 2D projection of RPS of rhythms as features. The genetic algorithm is selected as a significant feature [16]. After that, these features are fed to the K-nearest neighbor (KNN) classifier in a tenfold cross-validation strategy for classification of S and SF EEG signals.
2. PROPOSED METHOD

In this paper, EEG signals are separated into the delta, theta, alpha, beta and gamma rhythms using EWT. Then 2D projections of rhythms are plotted by RPS and considered their shape patterns, significant features are computed. Feature vector arrays are selected by Genetic algorithm; Finally, the KNN algorithm is classified as the EEG signals in S and SF groups. Figure 1 shows the block diagram of the proposed method.

The subset of A and B were recorded from five healthy subjects in eyes opened and closed conditions respectively. The subset of C and D were recorded from five patients who had completely recovered from seizure control after surgery of epileptic locations. The subset E is composed of EEG signals with epileptic seizure activities that are observed in the epileptogenic zone. In this work, signals in C and D subsets are considered as SF EEG signals, and signals in the E subset are considered as S EEG signals. Figure 2 shows an S and SF signal.

2.1. Database Used

The proposed method has been evaluated by the benchmark dataset which is freely downloadable from the Bonn University website [17]. This database consists of 5 subsets called A, B, C, D, and E, that each subset has 100 EEG signal were sampled at a rate of 173.61 Hz. The duration of each EEG signal is 23.6 second, so has 4096 samples.

![Fig.1. Proposed method.](image-url)
2.2. Empirical Wavelet Transform (EWT)

EWT decomposes signals by generating an adaptive filter bank corresponding to the input signal spectrum. Bandpass of the adaptive filter bank is determinate using proper segmentation of the spectrum [13, 14]. In the EWT toolbox, many methods are proposed to proper segmentation of the spectrum [18]. Segmentation of spectrum to [0-4 Hz], [4-8 Hz], [8-16 Hz], [16-30] and [30-60] bands will be resulted delta, theta, alpha, beta and gamma rhythms, respectively. For this propose, we set the cut-off frequencies as $f_{\omega} = \{4, 8, 16, 30, 60\}$ and use them for construct scaling function and wavelet functions. Filters of scaling function $\phi(\omega_f)$ and wavelet functions $\psi(\omega_f)$ are constructed in Fourier domain based on Littlewood-Paley and Meyer wavelets as follows [13]:

$$
\phi(\omega_f) = \begin{cases} 
1 & \text{if } |\omega_f| \leq (1-\lambda)f_i \\
\cos\left(\frac{\pi \beta(\lambda f_i)}{2}\right) & \text{if } (1-\lambda)f_i \leq |\omega_f| \leq (1+\lambda)f_i \\
0 & \text{otherwise}
\end{cases}
$$

(1)
\[
\psi_{i=1,2,...,5}(\omega_j) = \begin{cases} 
1 & \text{if } (1+\lambda)f_i \leq |\omega_j| \leq (1-\lambda)f_{i+1} \\
\cos\left(\frac{\pi\beta(\lambda f_{i+1})}{2}\right) & \text{if } (1-\lambda)f_{i+1} \leq |\omega_j| \leq (1-\lambda)f_{i+1} \\
\sin\left(\frac{\pi\beta(\lambda f_i)}{2}\right) & \text{if } (1+\lambda)f_i \leq |\omega_j| \leq (1+\lambda)f_i \\
0 & \text{otherwise}
\end{cases}
\]

where \( \beta(\lambda,\omega_j) = \beta\left(\frac{|\omega_j| - (1-\lambda)}{2\lambda \omega_j}\right) \)

Parameter defined as \( \lambda < \min\left(\frac{\omega_{i+1} - \omega_{i}}{\omega_{i+1} + \omega_{i}}\right) \),

make sure that the EWT coefficients are in \( L^2(\mathbb{R}) \) space.

In this work, \( \lambda \) parameter is computed to 0.1825. Also, \( \beta(y) \) is arbitrary function defined as:

\[
\beta(y) = \begin{cases} 
0 & \text{if } y \leq 0 \\
\beta(y) + \beta(1-y) = 1 & \forall y \in [0,1] \\
1 & \text{if } y \geq 1
\end{cases}
\]

Finally, each rhythm can be found by the inner product of EEG signals with corresponding filters.

The EWT filter bank generates a tight frame, transition band of the filters is very small, and pass-band and stop-band ripples are negligible in the band-pass filters [13, 14]. As a result, separated rhythms will have very little aliasing, which leads to more precise representation.

Figure 3 shows the EWT filter bank for EEG rhythms separation and Figure 4 shows the separated rhythms from S and SF signals. It should be noted that any frequency content greater than the highest frequency range of gamma is discarded as a noise signal.

2.3. The Reconstructed Phase Spaces

The reconstructed phase space (RPS) were used to show the nonlinear nature of the Stabilogram signal [19]. In this paper, RPS of rhythms is used as a visual image for the evaluation of the dynamical behavior of S and SF EEG signals. RPS has been used previously for the classification of normal and Attention deficit hyperactivity disorder.

![Fig. 3. EWT filter bank.](image)
(ADHD) EEG signals [20]. The RPS generation requires determination of delay time $\tau$ and embedding dimension $d$ which can be obtained by mutual information (MI) [14, 20] and the nearest neighbor (NN) method [14, 20], respectively. For the signal $V = \{v_1, v_2, v_3, ..., v_k\}$, where $K$ is the total number of data point, the RPS defined as:

$$Y_k = \left( V_k, V_{k+\tau}, ..., V_{k+(d-1)\tau} \right)$$  \hspace{1cm} (3)

where,

$$k = 1, 2, ..., K - (d - 1)\tau$$ \hspace{1cm} (4)

The computations for obtaining the $\tau$ and $d$ parameters are very heavy [14]. For this, in this work, $\tau$ and $d$ values are chosen empirically to 1 and 2, consecutively. The 2D projection of RPS of a signal obtained by plotting $V_k$ against $V_{i\tau}$, 2D RPS of rhythms of S and SF signals are show in figure 5.

2.4. Feature Extraction

2.4.1 Ellipse Area

It is clear form figure 5 that the 2D RPS of rhythms have elliptical patterns. It motivates us to compute the ellipse area of 2D RPS of rhythms for classification of S and SF EEG signals.

The procedure to calculate the ellipse area from the RPS can be given as [21].

Compute the mean values of $V_k$ and $V_{i\tau}$ as:

$$S_X = \sqrt{\frac{1}{K-1} \sum_{k=1}^{K-1} V_k^2}$$  \hspace{1cm} (5)
Fig. 5. From up to down shows the 2D RPS of the delta, theta, alpha, beta and gamma rhythms for a sample of SF (left) and S (right) EEG signal.
\[ S_Y = \sqrt{\frac{1}{K-1} \sum_{k=1}^{K-1} V_k^2} \]  
(6)

\[ S_{XY} = \frac{1}{K-1} \sum_{k=1}^{K-1} V_k V_{k+1} \]  
(7)

Compute C parameter as:

\[ C = \sqrt{\left( S_X^2 + S_Y^2 \right) - 4 \left( S_X^2 S_Y^2 - S_{XY}^2 \right)} \]  
(8)

\[ a = 1.7321 \sqrt{S_X^2 + S_Y^2 + C} \]  
(9)

\[ b = 1.7321 \sqrt{S_X^2 + S_Y^2 - C} \]  
(10)

From the parameters ‘a’ and ‘b’, the ellipse area is computed as equation (11):

\[ A = \pi ab \]  
(11)

2.4.2 Shortest Distance

S signals occupying more are in 2D RPS. In other words, scattering of 2D RPS of S signals are more than SF signals. For this reason, it seems that distance computation can be a useful feature.

If \( G(x_0, y_0) \) is a point on coordinate plane and \( ax + by + c = 0 \) is a line with slope \( m = -a/b \), the shortest distance \((shD)\) from point \( G \) to line \( ax + by + c = 0 \) is calculated as follow:

\[ shD = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}} \]  
(12)

The concept of these parameters is illustrated in figure 6.

For quantifying the scattering of data on coordinate plane, we compute the distance of \( G \) point to \( y = x \) and \( y = -x \) which known as 45 and 135 degree lines as follow:

\[ shD^{45} = \frac{|y_0 - x_0|}{\sqrt{2}} \]  
(13)

\[ shD^{135} = \frac{|x_0 + y_0|}{\sqrt{2}} \]  
(14)

Finally, for combination of these two distances as one parameter for quantifying the scattering of data, we computed the rectangle area which made by \( shD^{45} \) and \( shD^{135} \). In this work, summation of rectangle areas made by all points on RPS is computed as a feature. It can define as follow:

\[ shD = \sum_{i=1}^{n} (shD_i^{45} \times shD_i^{135}) \]  
(15)

where \( n \) is the number of points on RPS plane and \( i \) is the \( i^{th} \) point on RPS plane with \( G(x_i, y_i) \) coordinate.

2.5. Genetic Algorithm

Genetic algorithm has been proposed in 1996 as an optimization method in several applications like medical, engineering, physics, biology and statistics [16, 22]. Genetic algorithm is a search heuristic that is routinely used to generate useful solutions to
optimization and search problems [23]. It generates solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover [16, 23]. Genetic algorithms are one of the best ways to solve a problem for which little is known. Nowadays, the Genetic algorithm is one of the most used tools in machine learning applications as a feature selection technique. For the implementation of the Genetic algorithm in MATLAB, we used the written function by Jingwei Too [24]. It has 6 input parameters, namely: feat, label, N, T, CR and MR which are extracted features, labels, the number of chromosomes, the maximum number of generations, crossover rate, and mutation rate, respectively. Also, the outputs of MATLAB function are sFeat and Sf which indicate selected features and selected feature index, respectively. These parameters have been described in [16, 22, 23], for more information, please see these papers. We have selected the N, T, CR and MR parameter to 10, 100, 0.8 and 0.01, respectively. It should be noted that the KNN classification performance is used as a fitness function in Genetic algorithm algorithms.

2. 6. K Nearest Neighbor (KNN) Classifier

KNN is a supervised classifier with very easy theory and implementation. The KNN classifies any sample of test data considering their K closed neighbor samples in the train data [14]. Test samples belong to the group which has more members among K closed neighbors. The distance computation method and number of K are two parameters of the KNN classifier. In this work, City block distance is used by a varied number of k from 2 to 9 by step 1 to classify S and SF EEG signals. The sensitivity (SEN), specificity (SPE), accuracy (ACC), positive predictive value (PPV) and negative predictive value (NPV) parameters [22] are computed to evaluate the classifier performance. Figure 7 shows the binary KNN classification algorithm for classifying an input test sample considering to train samples with assuming that two features be extracted.

Fig. 7. Illustration of the KNN algorithm as a used classifier. The first, second, third and fourth circles around test data determine the one, seven, fourteen and twenty-one closed training data. The test data belongs to class A, B, A and B with assuming that k be 1, 7, 14 and 21, respectively.
3. RESULTS AND DISCUSSION

In this paper, we propose a method based on extracted rhythms in the EWT domain and RPS to the classification of S and SF EEG signals. EEG signals are decomposed to EEG rhythms using EWT. Then, 2D RPS of rhythms is plotted and A and shD are computed as features. The Kruskal–Wallis statistical test evaluated the features corresponding to their p-values [12, 14]. The lesser p-value indicates better discrimination between the S and SF classes. P-values of extracted features corresponding to each rhythm are written in Table 1.

It is evident that all rhythms show good discrimination between S and SF EEG signals (p ≈ 0). In other words, we could use from all of the computed features, but in order to reduce the complexity of the classifier, the Genetic algorithm selected the significant features. Table 2 gives the selected features (i.e. A and shD) from RPS of EEG signal rhythms (i.e. delta, theta, alpha, beta and gamma).

Significant features are fed to KNN classifier with these features is written in Table. 3.

It is clear from table 3 that selected features by KNN classifier can result in the highest classification ACC of 98.33 in S and SF classification tasks with a ten-fold cross-validation strategy. We have compared our proposed method with existing methods studied on the same database in Table 4. It is clearly evident that the proposed method archived the highest classification ACC. In [10], researchers have used EMD and SODP to detecting S and SF EEG signals. Although they reported the best accuracy of 97.75%, which are very close to the highest ACC

Table 1. Computed p-values for features.

<table>
<thead>
<tr>
<th>Rhythms</th>
<th>p-value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>shD</td>
</tr>
<tr>
<td>delta</td>
<td>$4.83 \times 10^{-25}$</td>
<td>$6.56 \times 10^{-29}$</td>
</tr>
<tr>
<td>theta</td>
<td>$3.31 \times 10^{-43}$</td>
<td>$2.15 \times 10^{-42}$</td>
</tr>
<tr>
<td>alpha</td>
<td>$5.64 \times 10^{-44}$</td>
<td>$9.80 \times 10^{-44}$</td>
</tr>
<tr>
<td>beta</td>
<td>$2.88 \times 10^{-44}$</td>
<td>$7.68 \times 10^{-44}$</td>
</tr>
<tr>
<td>gamma</td>
<td>$1.11 \times 10^{-40}$</td>
<td>$2.71 \times 10^{-42}$</td>
</tr>
</tbody>
</table>

Table 2. Selected features from EEG rhythms by BPSO which resulted to best performance in S and SF classification task.

<table>
<thead>
<tr>
<th>Rhythm</th>
<th>Extracted feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta</td>
<td>A</td>
</tr>
<tr>
<td>Theta</td>
<td>shD</td>
</tr>
<tr>
<td>Alpha</td>
<td>A</td>
</tr>
<tr>
<td>Beta</td>
<td>shD, A</td>
</tr>
<tr>
<td>Gamma</td>
<td>A</td>
</tr>
</tbody>
</table>
Table 3. The performance of KNN classifier with City block distance.

<table>
<thead>
<tr>
<th>Number of K</th>
<th>Classification Objective Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACC (%)</td>
</tr>
<tr>
<td>2</td>
<td>98.33</td>
</tr>
<tr>
<td>3</td>
<td>97.66</td>
</tr>
<tr>
<td>4</td>
<td>97.66</td>
</tr>
<tr>
<td>5</td>
<td>98</td>
</tr>
<tr>
<td>6</td>
<td>97.66</td>
</tr>
<tr>
<td>7</td>
<td>97.66</td>
</tr>
<tr>
<td>8</td>
<td>97.66</td>
</tr>
<tr>
<td>9</td>
<td>97.33</td>
</tr>
</tbody>
</table>

Table 4. Comparison of proposed method with the exiting work.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Classification task</th>
<th>Cross validation</th>
<th>ACC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[3]</td>
<td>N vs. S, F vs. S</td>
<td>Not used</td>
<td>88.83</td>
</tr>
<tr>
<td>[7]</td>
<td>N, F vs. S</td>
<td>Not used</td>
<td>95.33</td>
</tr>
<tr>
<td>[9]</td>
<td>N, F vs. S</td>
<td>Ten-fold</td>
<td>97.75</td>
</tr>
<tr>
<td>[12]</td>
<td>N, F vs. S</td>
<td>Ten-fold</td>
<td>97.5</td>
</tr>
<tr>
<td>Proposed method</td>
<td>N, F vs. S</td>
<td>Ten-fold</td>
<td>98.33</td>
</tr>
</tbody>
</table>

(98.33%) in our method, they are used EMD which suffers from the mode-mixing problem and noise-sensitive [9]. In even that circumstance, our proposed method is not noise sensitive.

4. CONCLUSION

Epileptic seizures in the human brain frequently manifest spikes in EEG signals [2] which can be analyzed visually by the experts. Visual inspection of long EEG
recording to detect the presence of epileptic seizures can be a cumbersome and time-consuming activity. This paper proposed a method based on rhythm separation using EWT for the classification of the S and SF EEG signals. A and shD are extracted as discrimination features from RPS of rhythms. Statistically, significant features are chosen by the Genetic algorithm and fed to KNN classifier in a ten-fold cross-validation strategy. The proposed method archived 98.33% classification ACC to detecting of S EEG signals using KNN classifier with City block distance. We have compared our proposed method with existing studies on the same database in Table 4. The proposed method archived the highest classification ACC compared with previous techniques. Our proposed method measured the variation and complexity of RPS of S and SF rhythms by A and shD features, respectively. These features could be a good parameter to discrimination between the S and SF classes because p-values were very near to zero. Recently, variation mode decomposition (VMD) has been propose instead of EWT method [25]. In future, the performance of the proposed feature with VMD will be evaluated.

REFERENCES


[10] R.B. Pachori, S. Patidar, "Epileptic seizure classification in EEG signals using second-order difference plot of


